CLAIMS

1.   1.   A storage system for use in a storage system cluster, the storage system compris-
2   ing:
3   a storage operating system including a cluster connection manager adapted to cre-
4   ate, destroy, and maintain one or more communication sessions with a cluster partner, the
5   cluster connection manager operatively interconnected with a set of cluster connection
6   manager clients.

1   2.   The storage system of claim 1 wherein one of the set of communication clients
2   comprises a failover monitor.

1   3.   The storage system of claim 1 wherein one of the set of cluster connection man-
2   ager clients comprises a non-volatile random access memory shadowing process.

1   4.   The storage system of claim 1 wherein the cluster connection manager is further
2   adapted to perform connection management operations in response to communications
3   from the connection manager clients.

1   5.   The storage system of claim 4 wherein the communications comprise an applica-
2   tion program interface function call.

1   6.   The storage system of claim 1 wherein the cluster connection manager is further
2   adapted to load balance the one or more communication sessions over a plurality of clus-
3   ter interconnect devices.

1   7.   The storage system of claim 1 wherein the cluster connection manager is further
2   adapted to perform a failover procedure for one or more communication sessions from a
3   failed cluster interconnect device to an operational cluster interconnect device.

1 8.   The storage system of claim 1 wherein the cluster connection manager is opera-

2 tively interconnected with a plurality of cluster interconnect devices.

1 9.   The storage system of claim 1 wherein the storage operating system comprises a

2 plurality of cluster connection managers.

1 10.   A storage operating system, executing on a storage system, the storage operating

2 system comprising:

3   a cluster connection manager adapted to manage a set of peer-to-peer connections

4 associated with a set of cluster connection manager clients executing on the storage sys-

5 tem.

1 11.   The storage operating system of claim 10 wherein the set of cluster connection

2 manager clients comprises a failover monitor.

1 12.   The storage operating system of claim 10 wherein the cluster connection manager

2 is further adapted to perform load balancing of the set of peer-to-peer connections over a

3 plurality of cluster interconnect devices.

1 13.   The storage operating system of claim 10 wherein the cluster connection manager

2 is further adapted to failover the set of peer-to-peer connections from a failed cluster in-

3 terconnect device to an operational cluster interconnect device.

1 14.   A method for initiating a peer-to-peer communication session, the method com-

2 prising the steps of:

3   creating an initial connection;

4   exchanging a set of peer connection information;

5   passing a set of client information to the cluster partner;

6   creating a set of appropriate communication ports;

7   alerting the cluster partner of a ready status; and

8   alerting a set of clients that the cluster partner is in a ready state.

1   15.   The method of claim 14 wherein the set of clients comprises a failover monitor

2   process.

1   16.   The method of claim 14 wherein the set of peer connection information comprises

2   a version number.

1   17.   The method of claim 14 wherein the step of passing a set of client information to

2   the cluster partner further comprises the steps of:

3   collecting, from a set of clients, the set of client information; and

4   transferring the collected set of client information to the cluster

1   18.   The method of claim 17 wherein the client information comprises a number of

2   communication ports required.

1   19.   The method of claim 17 wherein the set of client information further comprises an

2   amount of memory requested by a particular client.

1   20.   The method of claim 14 wherein the step of creating an initial connection further

2   comprises the step of using remote direct memory access primitives to create the initial

3   connection.

1   21.   The method of claim 14 wherein the step of creating an initial connection further

2   comprises the step of performing a series of remote direct memory access operations to

3   create the initial connection.

1   22.   A method for terminating a peer-to-peer communication session, the method

2   comprising the steps of:

3 alerting a set of clients of an impending termination of the communication ses-

4 sion;

5 closing, by the clients, a set of communication ports associated with the commu-

6 nication session; and

7 performing an initialization of a peer-to-peer communication session procedure.

1 23. The method of claim 22 wherein the set of communication ports comprises a set

2 of virtual interface connections.

1 24. The method of claim 22 wherein the set of clients comprises a failover monitor.

1 25. A storage operating system, executing on a storage system, the storage operating

2 system comprising:

3 a cluster connection manager having means to manage a set of peer-to-peer con-

4 nections associated with a set of cluster connection manager clients executing on the

5 storage system.

1 26. The storage operating system of claim 25 wherein the set of cluster connection

2 manager clients further comprises a failover monitor.

1 27. The storage operating system of claim 25 wherein the set of cluster connection

2 manager clients further comprises a nonvolatile random access memory shadowing proc-

3 ess.

1 28. A system configured to manage reliable peer communication among storage sys-

2 tems in a clustered environment, the system comprising:

3 one or more peer processes executing on each storage system partner; and

4 a cluster connection manager executing on each storage system partner, the clus-

5 ter connection manager creating a set of peer-to-peer connections between the one or

6 more peer processes executing on each storage system.

24